Hypothesis tests for Logistic Regression

$X_i \sim D$ indep; identical

$Y_i | X_i, \beta \sim Be\left(\frac{e^{X_i^T\beta}}{1+e^{X_i^T\beta}}\right)$

$l_n(\beta | Y, X) = \sum_{i=1}^{n} \left[Y_i X_i^T \beta - \log(1+e^{X_i^T\beta}) \right]$

$\hat\beta = \arg\max_\beta l_n(\beta | Y, X)$

<u>Goal:</u> Test $H_0: \beta_j = \beta_j^0$

$H_1: \beta_j \neq \beta_j^0$

① Wald test

② Power

③ Likelihood ratio test

$Y_i \in \{0, 1\}$, $X_i \in \mathbb{R}^p$, (X_i, Y_i) independent, $Y_i \sim Be(p_i)$

responses covariates

$$\mathbb{P}(Y_i = y_i | p_i) = \begin{cases} p_i & \text{if } y_i = 1 \\ (1-p_i) & \text{if } y_i = 0 \end{cases} = p_i^{y_i}(1-p_i)^{1-y_i}$$

$$= \exp\left(y_i \log p_i + (1-y_i)\log(1-p_i)\right) = \exp\Big(y_i \underbrace{\log\frac{p_i}{1-p_i}}_{=\eta_i = X_i^T\beta} + \log(1-p_i)\Big)$$

$\eta_i \in \mathbb{R}$, $\beta \in \mathbb{R}^p$

$$\Rightarrow \eta_i = \log\frac{p_i}{1-p_i} \Leftrightarrow e^{\eta_i} = \frac{p_i}{1-p_i} \Leftrightarrow (1-p_i)e^{\eta_i} = p_i$$

$$\Leftrightarrow p_i(1+e^{\eta_i}) = e^{\eta_i} \Leftrightarrow p_i = \frac{e^{\eta_i}}{1+e^{\eta_i}}$$

$$\Rightarrow \mathbb{P}(Y_i = y_i | X_i, \beta) = \exp\left(y_i X_i^T\beta + \log\left(1 - \frac{e^{X_i^T\beta}}{1+e^{X_i^T\beta}}\right)\right)$$

$$= \exp\left(y_i X_i^T\beta - \log\left(1+e^{X_i^T\beta}\right)\right)$$

$X_i \sim D$ indep; identical

$Y_i | X_i, \beta \sim Be\left(\frac{e^{X_i^T\beta}}{1+e^{X_i^T\beta}}\right)$

$\ell_n(\beta | Y, X) = \sum_{i=1}^{n}\left[Y_i X_i^T \beta - \log(1+e^{X_i^T\beta})\right]$

$\hat\beta = \underset{\beta}{\operatorname{argmax}}\, \ell_n(\beta | Y, X)$

Goal: Test $H_0: \beta_j = \beta_j^0$

$H_1: \beta_j \neq \beta_j^0$

1) Wald test

$T_n^W = n\frac{(\hat\beta_j - \beta_j^0)^2}{(\hat I(\hat\beta)^{-1})_{jj}}$, $\psi^W = \mathbb{1}\{T_n^W > q_\alpha(\chi_1^2)\}$

2) Power

3) Likelihood ratio test

1) $Y_1, \dots, Y_n \sim \mathbb{P}_\theta$, $\theta \in \mathbb{R}^d$, iid., $\hat\theta$ MLE, test $H_0: \theta = \theta^0$; $H_1: \theta \neq \theta^0$

$T_n = n \cdot (\hat\theta - \theta^0) \underbrace{I(\hat\theta)}_{= -\mathbb{E}[\nabla^2 \ell_1(\theta | Y_1)]} \cdot (\hat\theta - \theta^0) \xrightarrow[n\to\infty]{D} \chi_d^2$ under H_0

Theorem for MLE: $\sqrt{n}(\hat\theta - \theta) \xrightarrow[n\to\infty]{D} N(0, I(\theta)^{-1})$

$\Rightarrow n(\hat\theta - \theta) I(\theta)(\hat\theta - \theta) \xrightarrow[n\to\infty]{D} \chi_d^2$

$\underset{\text{Slutsky's Thm}}{\Rightarrow} n(\hat\theta - \theta) I(\hat\theta)(\hat\theta - \theta) \xrightarrow[n\to\infty]{D} \chi_d^2$, $\psi = \mathbb{1}\{T_n > q_\alpha(\chi_d^2)\}$

$1-\alpha$ qu.

Here: $\sqrt{n}(\hat\beta - \beta) \xrightarrow[n\to\infty]{D} N(0, I(\beta)^{-1}) \Rightarrow \sqrt{n}(\hat\beta_j - \beta_j) \xrightarrow[n\to\infty]{D} N(0, (I(\beta)^{-1})_{jj})$

$\hat I(\beta) = -\frac{1}{n}\sum_{i=1}^{n} \nabla^2 \ell_1(\beta | Y_i, X_i)$, use $\hat I(\hat\beta) \xrightarrow[n\to\infty]{\mathbb{P}} I(\beta)$

$\underset{\text{Slutsky}}{\Rightarrow} n\frac{(\hat\beta_j - \beta_j^0)^2}{(\hat I(\hat\beta)^{-1})_{jj}} \xrightarrow[n\to\infty]{D} \chi_1^2$

$X_i \sim D$ indep., identical

$Y_i | X_i, \beta \sim Be\left(\frac{e^{X_i^T\beta}}{1+e^{X_i^T\beta}}\right)$

$$l_n(\beta | Y, X) = \sum_{i=1}^{n} \left[Y_i X_i^T \beta - \log(1+e^{X_i^T\beta}) \right]$$

$\hat{\beta} = \arg\max_{\beta} l_n(\beta | Y, X)$

Goal: Test $H_0: \beta_j = \beta_j^0$

$H_1: \beta_j \neq \beta_j^0$

① Wald test

$$T_n^W = n \frac{(\hat{\beta}_j - \beta_j^0)^2}{(\hat{I}(\hat{\beta})^{-1})_{jj}}, \quad \psi = \mathbb{1}\{T_n^W > q_\alpha(\chi_1^2)\}$$

② Power

③ Likelihood ratio test

① ctd: $I(\beta) = -\mathbb{E}[\nabla^2 l_1(\beta | Y_1, X_1)]$

$$\hat{I}(\beta) = -\frac{1}{n}\sum_{i=1}^{n} \nabla^2 l_1(\beta | Y_i, X_i)$$

$$\partial_k l_1 = Y_i X_{ik} - \frac{1}{1+e^{X_i^T\beta}} \cdot e^{X_i^T\beta} X_{ik}$$

$$\partial_l \partial_k l_1 = -X_{ik} \frac{X_{il} e^{X_i^T\beta} \cdot (1+e^{X_i^T\beta}) - X_{il} e^{X_i^T\beta} \cdot e^{X_i^T\beta}}{(1+e^{X_i^T\beta})^2}$$

$$= -X_{ik} X_{il} \cdot \underbrace{\frac{e^{X_i^T\beta}}{(1+e^{X_i^T\beta})^2}}_{p_i(1-p_i)}$$

$$\hat{I}(\beta) = \frac{1}{n}\sum_{i=1}^{n} X_{ik} X_{il} \frac{e^{X_i^T\beta}}{(1+e^{X_i^T\beta})^2}$$

$X_i \sim D$ indep., identical

$Y_i \mid X_i, \beta \sim Be\left(\frac{e^{X_i^T\beta}}{1+e^{X_i^T\beta}}\right)$

$\ell_n(\beta \mid Y, X) = \sum_{i=1}^{n}\left[Y_i X_i^T\beta - \log\left(1+e^{X_i^T\beta}\right)\right]$

$\hat\beta = \arg\max_\beta \ell_n(\beta \mid Y, X)$

Goal: test $H_0: \beta_j = \beta_j^0$

$H_1: \beta_j \neq \beta_j^0$

1) Wald test

$T_n^W = n\frac{(\hat\beta_j - \beta_j^0)^2}{(\hat{I}(\hat\beta)^{-1})_{jj}}, \quad \psi = \mathbb{1}\{T_n^W > q_\alpha(\chi_1^2)\}$

$\hat{I}(\beta)_{kl} = \frac{1}{n}\sum_{i=1}^{n} X_{ik}X_{il}\frac{e^{X_i^T\beta}}{(1+e^{X_i^T\beta})^2}$

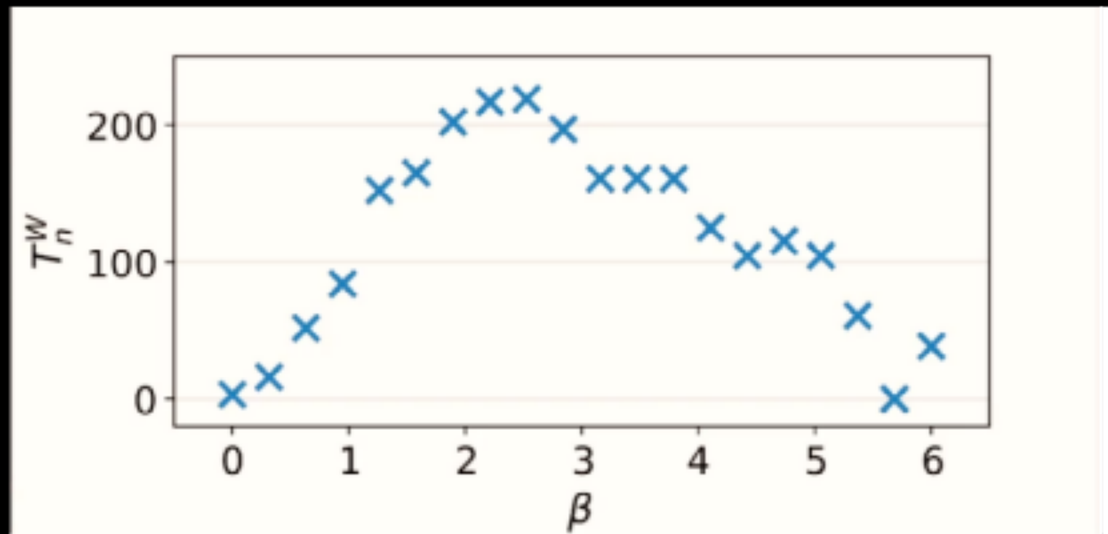
2) Power

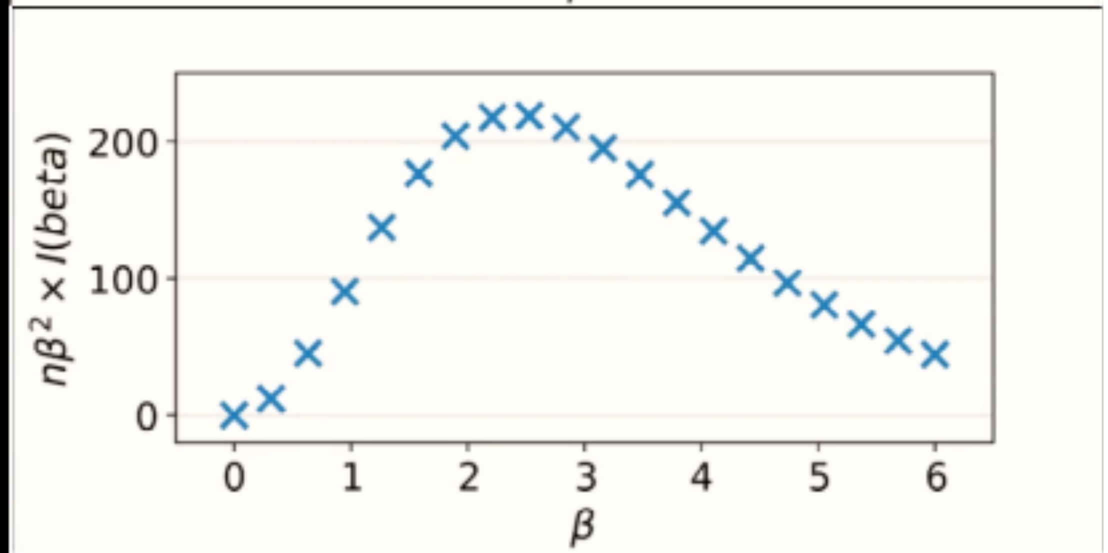
3) Likelihood ratio test

② $X_i \sim Be(\frac{1}{2})$, $Y_i \sim Be\left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right)$, $\beta \neq 0$, $\beta^0 = 0$

$\sqrt{n}\sqrt{\hat{I}(\hat\beta)}(\hat\beta - \beta + \beta) \text{ "}\sim\text{" } N(\sqrt{nI(\beta)}\,\beta, 1)$

$T_n^W = n\hat\beta^2\hat{I}(\hat\beta) \text{ "}\sim\text{" } \chi_1^2(nI(\beta)\beta^2)$ non-central χ_1^2





$X_i \sim D$ indep., identical

$Y_i | X_i, \beta \sim Be\left(\frac{e^{X_i^T\beta}}{1+e^{X_i^T\beta}}\right)$

$l_n(\beta|Y,X) = \sum_{i=1}^{n}\left[Y_i X_i^T \beta - \log(1+e^{X_i^T\beta})\right]$

$\hat\beta = \arg\max_\beta l_n(\beta|Y,X)$

Goal: Test $H_0: \beta_j = \beta_j^0$

$H_1: \beta_j \neq \beta_j^0$

① Wald test

$T_n^W = n\frac{(\hat\beta_j - \beta_j^0)^2}{(\hat I(\hat\beta)^{-1})_{jj}}$, $\psi = \mathbb{1}\{T_n^W > q_\alpha(\chi_1^2)\}$

$\hat I(\beta)_{kl} = \frac{1}{n}\sum_{i=1}^{n} X_{ik}X_{il}\frac{e^{X_i^T\beta}}{(1+e^{X_i^T\beta})^2}$

② Power

③ Likelihood ratio test

② $X_i \sim Be(\frac{1}{2})$, $Y_i \sim Be\left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right)$, $\beta \neq 0$, $\beta^0 = 0$

$$\sqrt{n}\sqrt{\hat I(\hat\beta)}(\hat\beta - \beta + \beta) \text{ "}\sim\text{" } N(\sqrt{n I(\beta)}\,\beta, 1)$$

$$T_n^W = n\hat\beta^2 \hat I(\hat\beta) \text{ "}\sim\text{" } \chi_1^2(nI(\beta)\beta^2) \quad \text{non-central } \chi_1^2$$

$$I(\beta) = \mathbb{E}\left[X_i^2\frac{e^{X_i\beta}}{(1+e^{X_i\beta})^2}\right] = \frac{1}{2}\cdot\frac{e^\beta}{(1+e^\beta)^2}$$

$$\beta^2 I(\beta) = \frac{1}{2}\cdot\frac{\beta^2 e^\beta}{(1+e^\beta)^2} \underset{*}{\sim} \frac{1}{2}\frac{2\beta\cancel{e^\beta} + \beta^2\cancel{e^\beta}}{2(1+e^\beta)\cancel{e^\beta}}$$

$$= \frac{1}{4}\frac{2\beta+\beta^2}{1+e^\beta} \underset{*}{\sim} \frac{1}{4}\frac{2+2\beta}{e^\beta} \underset{*}{\sim} \frac{1}{2}\cdot\frac{1}{e^\beta} \xrightarrow[\beta\to+\infty]{} 0$$

L'Hospital's rule: *

$f(x), g(x)$, $x \to c$

$f(x) \to \infty$, $g(x) \to \infty$, f, g differentiable

$\Rightarrow \lim_{x\to c}\frac{f(x)}{g(x)} = \lim_{x\to c}\frac{f'(x)}{g'(x)}$

$X_i \sim D$ indep.; identical

$$Y_i | X_i, \beta \sim Be\left(\frac{e^{X_i^T\beta}}{1+e^{X_i^T\beta}}\right)$$

$$l_n(\beta|Y,X) = \sum_{i=1}^{n}\left[Y_i X_i^T \beta - \log\left(1+e^{X_i^T\beta}\right)\right]$$

$$\hat\beta = \arg\max_{\beta} l_n(\beta|Y,X)$$

Goal: Test $H_0: \beta_j = \beta_j^0$

$H_1: \beta_j \neq \beta_j^0$

① Wald test

$$T_n^W = n\frac{(\hat\beta_j - \beta_j^0)^2}{(\hat{I}(\hat\beta)^{-1})_{jj}}, \quad \psi = \mathbb{1}\{T_n^W > q_\alpha(\chi_1^2)\}$$

$$\hat{I}(\beta)_{kl} = \frac{1}{n}\sum_{i=1}^{n} X_{ik}X_{il}\frac{e^{X_i^T\beta}}{(1+e^{X_i^T\beta})^2}$$

② Power

③ Likelihood ratio test

③ Likelihood ratio test: calculate $\hat\beta$ &

$$\hat\beta^c = \arg\max_{\beta_j = \beta_j^0} l_n(\beta|Y,X)$$

$$T_n^{LR} = 2\cdot\left(l_n(\hat\beta|Y,X) - l_n(\hat\beta^c|Y,X)\right) \xrightarrow[n\to\infty]{D} \chi^2_{p-(p-1)} = \chi^2_1$$

$$\psi^{LR} = \mathbb{1}\{T_n^{LR} > q_\alpha(\chi_1^2)\}$$